

枣树叶面积估算模型构建

张萌, 张德安, 鲁晓燕, 杨伟伟*

(石河子大学农学院·特色果蔬栽培生理与种质资源利用兵团重点实验室, 新疆石河子 832000)

摘要:【目的】构建并评估多个枣树叶面积估算模型, 筛选出适宜的枣树叶面积无损测量模型。【方法】以叶片形态差异显著的8个枣树品种为试材, 通过数字扫描仪获取叶片长度(LL)、宽度(LW)、叶面积(LA)等形态参数, 基于决定系数(R^2)、均方根误差(RMSE)、AIC(Akaike Information Criterion, 赤池信息量准则,)对建立的17个线性及非线性叶面积模型进行筛选及适宜性评价。【结果】共获得3 287枚叶片形态参数, 各叶片形态参数间差异显著, 尤其LA变异系数最大, 达64.44%。以LL和LW复合变量为自变量的线性模型5 $LA = a(LL \times LW)$ 、模型6 $LA = a(LL+LW)^2$ 、模型9 $LA = a LL^2 + b LW^2$ 和非线性模型17 $LA = (LL \times LW)^b$ 的精度可满足各个品种叶面积的估算, 但需要根据品种拟合模型系数。基于全部数据的总体模型5和6的精度可用于各个枣品种叶面积估算, 模型系数分别为0.703 5和0.164 6, 确定系数分别为99.40%和98.84%, RMSE分别为0.46 cm^2 和0.64 cm^2 , 两模型参数的准确估计需要的样本数量应大于250。【结论】筛选出的枣树单叶面积模型精度可满足供试各品种叶片面积估算的要求, 无需针对单个品种独立建模。

关键词: 枣树; 叶面积; 模型; AIC

中图分类号: S665.1

文献标志码: A

文章编号: 1009-9980(2020)12-1964-10

Establishment of leaf area estimation model for jujube trees

ZHANG Meng, ZHANG De'an, LU Xiaoyan, YANG Weiwei*

(College of Agriculture, The Key Laboratory of Special Fruits and Vegetables Cultivation Physiology and Germplasm Resources Utilization in Xinjiang Production and Construction Group, Shihezi University, Shihezi 832000, Xinjiang, China)

Abstract: 【Objective】The study aims to establish a reliable and accurate leaf area (LA) estimation model based on leaf length (LL) and leaf width (LW) which can be measured non-destructively. Leaf area is usually determined by destructive methods, which are time-consuming and impossible to make successive measurements on the same leaf samples. Non-destructive portable leaf area laser scanners are expensive and complex. The model-based leaf area estimation overcomes the above defects. However, the optimal model needs to be selected from candidate models for jujube trees. 【Methods】Eight jujube cultivars with significant differences in leaf morphology, including 'Junzao' 'Lizao' 'Qiyuexian' 'Suanzao' 'Zanhuangdazao' 'Fengmiguan' 'Gagazao' and 'Jinsixiaozao', were used as the materials. Leaves were sampled and then scanned with a laser scanner to obtain digital images. Leaf morphological parameters, such as LL, LW, LA and petiole length (PL), were extracted using ImageJ software. There were 17 candidate models established and compared. Each model was established and validated using pooled data from all cultivars and separately for each cultivar. A further validation was conducted for the general models constructed with pooled data to evaluate their applicability to different individual cultivars. The root mean square error (RMSE), determination coefficient (R^2) and Akaike information criterion (AIC) were used to evaluate the accuracy of the established models for leaf area through the comparisons between predicted values and actually measured values of leaf area. Each dataset was separated randomly into training dataset (75% of the data) and testing dataset (25% of the data) to fit and validate models by using R software. 【Results】3 287 leaves were sampled and significant differences

收稿日期: 2020-06-29 接受日期: 2020-08-30

基金项目: 石河子大学高层次人才科研启动金(RCSX201726); 特色果蔬栽培生理与种质资源利用兵团重点实验室开放课题(002)

作者简介: 张萌, 男, 在读硕士研究生, 研究方向: 果树生理生态。Tel: 18899598823, E-mail: 463768867@qq.com

*通信作者 Author for correspondence. Tel: 18899592799, E-mail: yww2567@126.com

were found among the cultivars in LL, LW, LA, PL, LL/LW ratio, PL/(PL+LL) ratio and roundness. The coefficient of variation (*CV*) in leaf area (LA) was the largest (64.44%), and followed by PL (43.41%), LW (37.24%) and LL (32.20%). LL, LW and their integrated variables, such as LL+LW, LL×LW, LL² and LW² etc., could be used as the independent variables to predict the LA for all cultivars. By using all data together, the 17 models showed varied predictivity ($71.66 \leq R^2 \leq 99.52\%$, $0.51 \leq \text{RMSE} \leq 3.09 \text{ cm}^2$ and $3\ 769.08 \leq \text{AIC} \leq 16\ 947.72$). Among those models, model 16 $\text{LA} = a (\text{LL} \times \text{LW})^b$ had the highest accuracy ($R^2=99.52\%$, $\text{RMSE}=0.44 \text{ cm}^2$, $\text{AIC}=3\ 769.08$), followed in descending order by model 5 $\text{LA} = a (\text{LL} \times \text{LW})$ ($R^2=99.40\%$, $\text{RMSE}=0.46 \text{ cm}^2$, $\text{AIC}=4\ 076.65$), model 9 $\text{LA} = a \text{LL}^2 + b \text{LW}^2$ ($R^2=99.34\%$, $\text{RMSE}=0.51 \text{ cm}^2$, $\text{AIC}=4\ 487.99$), model 14 $\text{LA} = a (\text{LL} + \text{LW})^b$ ($R^2=98.87\%$, $\text{RMSE}=0.65 \text{ cm}^2$, $\text{AIC}=6\ 469.08$), model 6 $\text{LA} = a (\text{LL} + \text{LW})^2$ ($R^2=98.84\%$, $\text{RMSE}=0.64 \text{ cm}^2$, $\text{AIC}=6\ 563.66$), and model 17 $\text{LA} = (\text{LL} \times \text{LW})^b$ ($R^2=98.72\%$, $\text{RMSE}=0.70 \text{ cm}^2$, $\text{AIC}=6\ 627.65$). Estimation with all the models for each cultivar showed that models 5, 6, 9 and 17 met the requirements of accuracy to predict the leaf area for all the cultivars. The coefficient of models 5, 6 and 17, had smallest *CV* among cultivars (2.71%, 2.91% and 3.41%, respectively) as compared to the other models. When the all-data models were validated by individual cultivars, models 5 and 6 still had a high accuracy of leaf area estimation, with the lowest R^2 appearing in ‘Suanzao’ in model 6, which was 96.80%, and the largest decrease in RMSE appearing in ‘Lizao’ in the same model, which was 0.145 1 cm^2 . Yet, when compared the R^2 between cultivar specified and all-data based model 17, the R^2 in cultivar specified model 17 for ‘Suanzao’ and ‘Jinsixiaozao’ decreased from 97.27% and 94.22% to 84.16% and 83.51% in all-data model, respectively. The model coefficients for model 5 and 6 were 0.703 5 and 0.164 6, respectively. LA was estimated with a higher accuracy by employing LW alone compared to LL alone, irrespective of cultivars. LA estimation was not always improved by employing both LL and LW as compared to employing single variable. LA can be estimated by employing LW alone with model $\text{LA} = a \text{LW}^b$, but the model coefficients need to be adjusted according cultivar. The minimum number of leaves was 250 to estimate the model coefficients for models 5 and 6 with relative lower errors. **【Conclusion】** For a specific cultivar, models 5, 6, 9 and 17 can meet the requirements of accuracy to predict the leaf area without the use of any expensive instruments but need to be fitted to obtain specific model coefficients for individual cultivars. Among the above 4 models, models 5 and 6 can be used to estimate leaf area accurately using the same model coefficients, irrespective of cultivar. The minimum sample size of 250 is required for models 5 and 6 to obtain reliable model coefficients. The accuracy leaf area model will be helpful to develop three-dimensional virtual jujube trees with accurate leaf dimensions.

Key words: Jujube trees; Leaf area; Modeling; AIC

叶片是植物最重要的源器官。叶片面积对果树光能利用、光合、呼吸、养分和水分利用、病虫害防治、花芽分化、坐果及果实品质和产量的形成具有显著影响^[1-5],是果树生理生化、遗传育种及栽培的重要研究指标^[6-8],也是果树冠层相关模型建立的基础之一^[9-10]。因此,获取精确的叶面积对于果树合理密植、冠层及群体结构调整、病虫害准确预测等具有重要意义。

叶面积的获得常采取破坏性测量方法,如打孔法、方格法、称重法和图像法,费时费力且无法实现动态测定^[7]。叶面积仪可动态测定,但仪器昂贵。叶

面积模型估算利用易获取的易测因子推算叶片面积,通过建立叶片长度、宽度或长和宽的复合变量与叶面积间的线性或非线性回归关系,进行叶面积的估算。叶面积估算模型已广泛应用于樱桃、葡萄、猕猴桃、桃、梨等果树^[1,7,11-15],是一种高效的叶面积获取方法。

枣是我国最古老,也是改革开放以来发展最快的果树之一,是我国干果中产量第一的主栽果树^[16]。杨伟等^[8]通过直接测定枣叶面积和产量,认为枣的产量与叶面积直接相关,密植枣树适宜的叶面积指数为 $2.33 \text{ m}^2 \cdot \text{m}^{-2}$ 。研究人员曾通过测定枣树叶

片面积与叶片长度、宽度,并建立回归关系进行叶片面积的估算^[5,8,17]。然而,有多种基于叶长、叶宽及二者复合变量的叶面积估算模型,枣树的适宜模型还未进行过筛选。同其他果树一样,枣树叶面积估算模型的建立主要针对特定品种,每个品种需要建立一个特异的叶面积估算模型^[5],同一个模型是否适用于不同品种还未有报道。同时,模型参数的拟合取决于采集的样本数据量,样本数据量会直接影响模型参数灵敏度,枣树叶面积估算需要的样本数据量,需要进一步深入研究。

笔者以叶片形态差异显著的8个枣树品种为试材,通过图像分析获取叶片形态参数,如叶片长度、叶片宽度、叶面积、叶柄长度等。首先,对不同品种间叶片形态参数进行比较分析,明确品种间叶片形态差异。其次,根据文献及本研究构建的17个常用线性和非线性叶面积估算模型,分别进行单个品种及不考虑品种差异的总体样本模型的拟合与验证,筛选适宜枣树叶面积的估算模型,并对总体模型的适宜性进行评价。通过随机设定样本数量,进一步明确枣树叶面积模型构建所需要的样本数量。枣树叶面积的精确估算,可为探讨枣树生长发育和外界环境条件的关系,研究枣树生理生态和遗传特性,以及其他植物叶面积模型构建和筛选提供数据、理论基础及方法借鉴。

1 材料和方法

1.1 试验材料

试验于2017年8月在新疆农垦科学院枣资源圃进行,地处天山北麓中段,属典型的温带大陆性气候,冬季长而严寒,夏季短而炎热,昼夜温差大,降水量少。该地区无霜期为168~171 d,≥10℃的活动积温3 570~3 729℃,年平均气温7.5~8.2℃,年降水量125.0~207.7 mm。日照充沛,年日照时数2 721~2 818 h。选取资源圃‘骏枣’‘梨枣’‘七月鲜’‘酸枣’‘赞皇大枣’‘蜂蜜罐’‘嘎嘎枣’‘金丝小枣’为试验材料。所有枣树皆为1 a(年)生植株,砧木为酸枣,春季平茬,枣树长出7~9个二次枝时打顶,株行距1 m×1 m,灌溉采用沟灌方法,土壤为壤土,行间未覆草,田间整体管理水平高。

1.2 试验设计与方法

1.2.1 叶片形态数据的测定 于2017年8月15—17日随机采集各品种枣树上层、中层、下层完整且

无病害的枣吊,每个品种取枣吊15~30个。为防止叶片失水变形,所有枣吊带叶片采集后装入避光冰盒,并带回实验室,使用Epson(V330)扫描仪扫描所有叶片,获得叶片数字图像。叶片采集后3 h内完成所有叶片的扫描工作。使用ImageJ软件计算叶片的形态指标,包括叶柄长度(petiole length, PL)、叶片长度(leaf length, LL)、叶片宽度(leaf width, LW)和叶面积(leaf area, LA),基于测定数据,计算叶片长宽比、叶柄/叶长比及圆度。本试验总计获取3 287枚叶片的形态参数,由于各品种枣吊长度和节间长度不一,各品种叶片数量范围介于303~453。

1.2.2 叶片面积模型构建、筛选及验证 植物叶面积的估算主要基于叶片形态特征参数。因此,参考已发表文献并结合本试验特点,本研究共构建了17个枣树叶面积估算的目标模型(表1),包括9个一元或二元线性模型(模型1~9)以及8个非线性模型(10~17)。

表1 枣树叶面积估算候选模型

Table 1 Candidate models for estimation of leaf area in jujube trees

模型类型 Model type	序号 Order	变量 Variable	模型 Model
线性模型 Linear model	1	LL	LA = a LL
	2	LW	LA = a LW
	3	LL ²	LA = a LL ²
	4	LW ²	LA = a LW ²
	5	LL×LW	LA = a (LL×LW)
	6	LL+LW	LA = a (LL+LW) ²
	7	LL+LW	LA = a (LL+LW) ³
	8	LL, LW	LA = a LL + b LW
	9	LL ² , LW ²	LA = a LL ² + b LW ²
非线性模型 Non-linear model	10	LL	LA = a LL ^b
	11	LL	LA = LL ^b
	12	LW	LA = a LW ^b
	13	LW	LA = LW ^b
	14	LL+LW	LA = a (LL+LW) ^b
	15	LL+LW	LA = (LL+LW) ^b
	16	LL×LW	LA = a (LL×LW) ^b
	17	LL×LW	LA = (LL×LW) ^b

利用各个模型分别对所有数据和各个品种的数据进行拟合,获取模型参数,构建不区分品种的总体模型和各个品种的特异模型。模型拟合时,随机选取各类样本总量的75%和25%作为训练集和测试集,利用训练集拟合模型。线性及非线性模型的拟

合使用 R 的 ‘lm’ 和 ‘nls’ 命令^[18]。利用测试集对模型精度进行验证,比较观测值与模拟值,评价指标包括决定系数(R^2)、均方根误差(Root Mean Square Error, RMSE)和 AIC (Akaike information criterion, 赤池信息量准则)^[19],各参数计算公式如下:

$$R^2 = 1 - \frac{\sum_{i=1}^n (AS_i - MS_i)^2}{\sum_{i=1}^n (AS_i - \overline{AS_i})^2}; \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (AS_i - MS_i)^2}{n}}; \quad (2)$$

$$AIC = -2 * \log \text{likelihood} + k * n. \quad (3)$$

其中, AS_i 为实测值, $\overline{AS_i}$ 为观测平均值, MS_i 为模型模拟值, n 为样本数, $\log \text{likelihood}$ 为模型最大似然估计, $k=2$ 为常数。

应用 R 软件制图及其他数据的统计分析^[18]。

2 结果与分析

2.1 不同品种叶片形态参数差异分析

根据测定的叶片形态参数(表 2),各个品种间各叶片形态参数均具有显著差异($p < 0.001$)。各形态参数具有很大的变异度,其中叶面积和叶柄长度的最大值分别是最小值的 70 和 21 倍。除圆度外,各叶片形态参数最大值与平均值的比值均大于 2,其中叶面积的比值最大,为 3.91。叶柄长度、叶片长度、叶面积及叶柄长度/(叶柄长度+叶片长度)的最大值均出现于骏枣品种。叶片圆度的变异系数最低(17.53%),而叶面积的变异系数最大(64.44%)。各品种的各叶片形态参数间差异极显著(表 3),叶片长度大的品种,其叶片宽度、叶柄长度、叶片面积、叶片长度/宽度比、圆度和叶柄比也大于其他大的品种。‘骏枣’‘嘎嘎枣’‘蜂蜜罐’‘梨枣’和‘赞皇大枣’的各叶片形态参数平均值大于或等于总体样本的均值,而‘七月鲜’‘金丝小枣’和‘酸枣’则小于总样本

表 2 枣树叶片形态特征参数统计值

Table 2 Leaf morphological parameters for the 8 jujube cultivars

参数 Parameter	中位数 Median value	最大值 Maximum value	最小值 Minimum value	平均值 Average value	变异系数 CV/%
叶片长度 Leaf length/cm	4.34	11.03(骏枣 Junzao)	1.03(梨枣 Lizao)	4.46	32.20
叶片宽度 Leaf width/cm	2.56	5.86(蜂蜜罐 Fengmiguan)	0.62(酸枣 Suanzao)	2.67	37.24
叶柄长度 Petiole length/cm	0.70	2.86(骏枣 Junzao)	0.14(酸枣 Suanzao)	0.76	43.41
叶片面积 Leaf area/cm ²	7.86	36.56(骏枣 Junzao)	0.52(七月鲜 Qiyuexian)	9.35	64.44
叶片长宽比 L/W ratio	1.65	3.45(酸枣 Suanzao)	0.82(蜂蜜罐 Fengmiguan)	1.72	16.96
圆度 Roundness	0.13	0.25(酸枣 Suanzao)	0.06(酸枣 Suanzao)	0.13	17.53
叶柄长度/(叶柄长度+叶片长度) PL/(PL+LL)	0.13	0.40(骏枣 Junzao)	0.05(蜂蜜罐 Fengmiguan)	0.14	25.58

表 3 不同枣树品种形态特征参数平均值

Table 3 Mean values of leaf morphological parameters for each jujube cultivar

品种 Cultivar	叶片长度 Leaf length/cm	叶片宽度 Leaf width/cm	叶柄长度 Petiole length/cm	叶面积 Leaf area/ cm ²	叶片长度与 叶片宽度比 LL/LW	圆度 Roundness	叶柄长度与(叶柄 长度+叶片长度)比 PL/(PL+LL)	叶片形状 Leaf shape
骏枣 Junzao	6.04 a	3.77 a	1.14 a	15.75 a	2.22 a	0.16 a	0.16 a	长卵形 Long ovate
嘎嘎枣 Gagazao	5.43 b	3.60 b	1.03 b	14.76 b	1.95 b	0.15 b	0.16 a	长卵形 Long ovate
蜂蜜罐 Fengmiguan	5.39 b	3.55 b	0.89 c	14.50 b	1.77 c	0.14 c	0.16 a	卵状披针形 Ovate-lanceolate
梨枣 Lizao	4.39 c	2.82 c	0.70 d	9.41 c	1.72 d	0.13 d	0.14 b	卵状披针形 Ovate-lanceolate
赞皇大枣 Zanhuangdazao	4.39 c	2.39 d	0.69 d	7.81 d	1.71 d	0.13 e	0.14 b	卵状披针形 Ovate-lanceolate
七月鲜 Qiyuexian	4.13 d	2.29 e	0.69 d	7.28 d	1.57 e	0.13 e	0.14 bc	卵状披针形 Ovate-lanceolate
金丝小枣 Jinsixiaozao	3.81 e	2.15 f	0.52 e	5.92 e	1.50 f	0.12 f	0.13 c	宽卵圆形 Wide ovoid
酸枣 Suanzao	2.69 f	1.23 g	0.42 f	2.55 f	1.44 g	0.10 g	0.11 d	卵圆形 Oval
p 值 p value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	-

注: p 值为品种间方差分析结果,同一列不同小写字母表示在 $p < 0.05$ 水平差异显著。

Note: P values represents the results from analysis of variance among cultivars, different small letters within a column are significantly different at $p < 0.05$.

的均值。‘骏枣’单个叶片平均面积最大(表3),为15.75 cm²,其余为‘嘎嘎枣’>‘蜂蜜罐’>‘梨枣’>‘赞皇大枣’>‘七月鲜’>‘金丝小枣’>‘酸枣’。

2.2 叶面积总体模型验证及筛选

不考虑品种差异,17个模型的拟合系数及模拟精度见表4,各个模型模拟叶面积的精度存在明显

表4 枣树叶片面积与叶片形状属性之间的估算模型的误差统计量

Table 4 Error statistics of the estimated models between jujube leaf area and jujube leaf shape attributes

模型序号 Model No.	模型系数a Model coefficient a	模型系数b Model coefficient b	决定系数 Determination coefficient, R ² /%	均方根误差 Root mean square error, RMSE/cm ²	AIC信息准则 Akaike information criterion, AIC
1	2.252 5		71.66	3.09	16 947.72
2	3.781 3		81.07	2.60	15 983.27
3	0.420 8		91.35	1.75	13 027.25
4	1.103 7		93.15	1.59	12 350.10
5	0.703 6		99.40	0.46	4 076.65
6	0.165 0		98.84	0.64	6 563.66
7	0.015 8		79.12	2.79	15 912.22
8	0.081 7	3.649 2	78.63	2.81	15 984.81
9	0.203 1	0.595 2	99.34	0.51	4 487.99
10	0.528 4	1.872 8	91.61	1.68	12 883.23
11	1.513 8		88.53	2.01	14 021.74
12	1.520 7	1.760 1	95.08	1.38	11 609.64
13	2.062 9		91.66	1.73	12 726.37
14	0.181 8	1.957 0	98.87	0.65	6 469.08
15	1.187 2		84.33	2.53	15 046.14
16	0.763 0	0.973 3	99.52	0.44	3 769.08
17	0.887 0		98.72	0.70	6 627.65

的差异(图1)。各模型R²、RMSE和AIC值分别为71.00%~99.52%、0.51~3.09 cm²和3 769.08~16 947.72。模型5、6、9、14、16和17的模拟值和观测值分布于1:1线附近,具有良好的预测能力,其R²值均大于98%,其中模型5、9和16的R²值大于99%。模型1、2、8和15模拟较小叶片时,易高估叶片面积;而模拟较大叶片面积时,易低估叶片面积,模型7则正好相反。模型3、4、10和13可以准确模拟较小叶片的面积,但随叶片面积增加,模拟值随机偏大或偏小。模型11和12虽然也可以模拟较小叶片面积,但易低估较大叶片叶面积。所有模型中,模型16的精度最高(R²=99.52%, RMSE=0.44 cm², AIC=3 769.08),其次为模型5(R²=99.40%, RMSE=0.46 cm², AIC=4 076.65)、模型9(R²=99.34%, RMSE=0.51 cm², AIC=4 487.99)、模型14(R²=98.87%, RMSE=0.65 cm², AIC=6 469.08)、模型6(R²=98.84%, RMSE=0.64 cm², AIC=6 563.66)、模型17(R²=98.72%, RMSE=0.70 cm², AIC=6 627.65)。

当模型变量包含叶片长度和宽度两个自变量时,模拟精度并未优于单个形态参数的模型(表4,

图1)。以单个叶片形态参数为变量的模型中,模型12表现最好,其R²=95.08%, RMSE = 1.38 cm², AIC= 11 609.64。

2.3 各品种叶面积模型验证及筛选

图2为各个品种分别拟合17个模型的模拟精度统计量。基于R²和RMSE值,对各个品种叶面积模拟效果最好的6个模型分别为模型5、6、9、14、16和17(94.09 ≤ R² ≤ 99.66, 0.09 ≤ RMSE ≤ 1.20 cm²)。基于AIC值,对各个品种叶面积模拟效果最好的6个模型分别为模型5、6、9、11、15和17(-611 ≤ AIC ≤ 1 211)。因此,不同枣品种使用模型5、6、9和17的精度可达到叶面积估算的要求。

从图3看出,各个品种拟合各个模型的模型参数变化幅度不一,变化幅度越小,说明模型参数在各个品种间的差异越小。对模型参数a而言,模型5、6和17的变异系数最小,分别为2.71%、2.91%和3.41%,且各个品种模型参数a与不考虑品种差异的总体模型参数a接近。模型1和10的模型参数a的变异系数可达31.16和37.05%。对于模型参数b,模型12、14和16的变异系数最小,分别为4.46%、

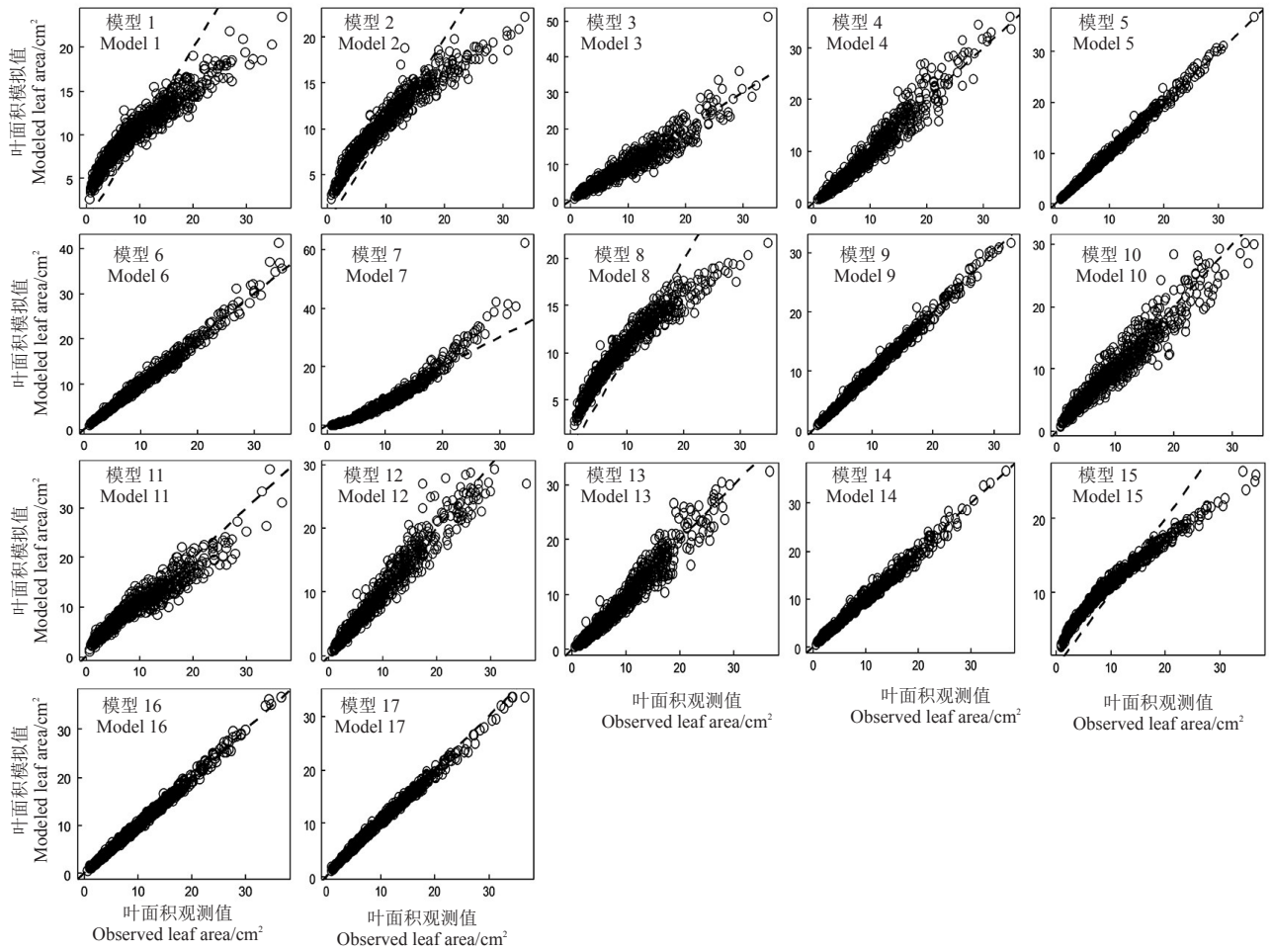
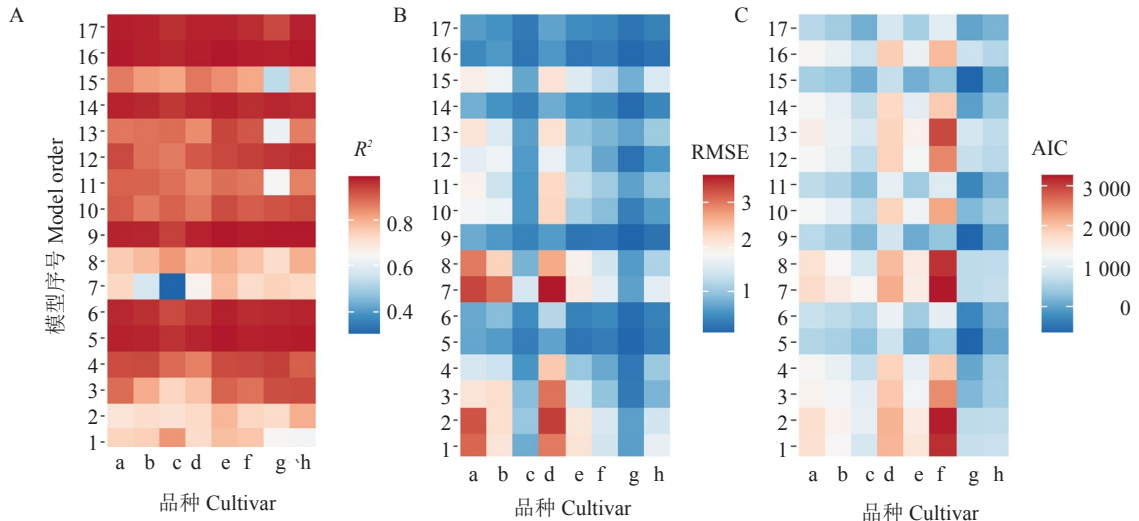


图 1 各总体模型观测值与模拟值散点图 (虚线为 1 : 1 线)

Fig. 1 Correlations between observed and predicted leaf area for each model by pooling all data together (dashed line means 1 : 1 line)



a. 蜂蜜罐; b. 嘎嘎枣; c. 金丝小枣; d. 骏枣; e. 梨枣; f. 七月鲜; g. 酸枣; h. 赞皇大枣。

a. Fengmiguan; b. Gagazao; c. Jinsixiaozao; d. Junzao; e. Lizao; f. Qiyuexian; g. Suanzao; h. Zhanhuangdazao.

图 2 各个品种不同模型模拟精度决定系数 (R^2 , A)、均方根误差 (RMSE, B) 和赤池信息量准则 (AIC, C) 热图

Fig. 2 Heatmaps for validation coefficients of determination coefficient (R^2 , A), root mean square error (RMSE, B) and Akaike information criterion (AIC, C) for each model in 8 jujube cultivars

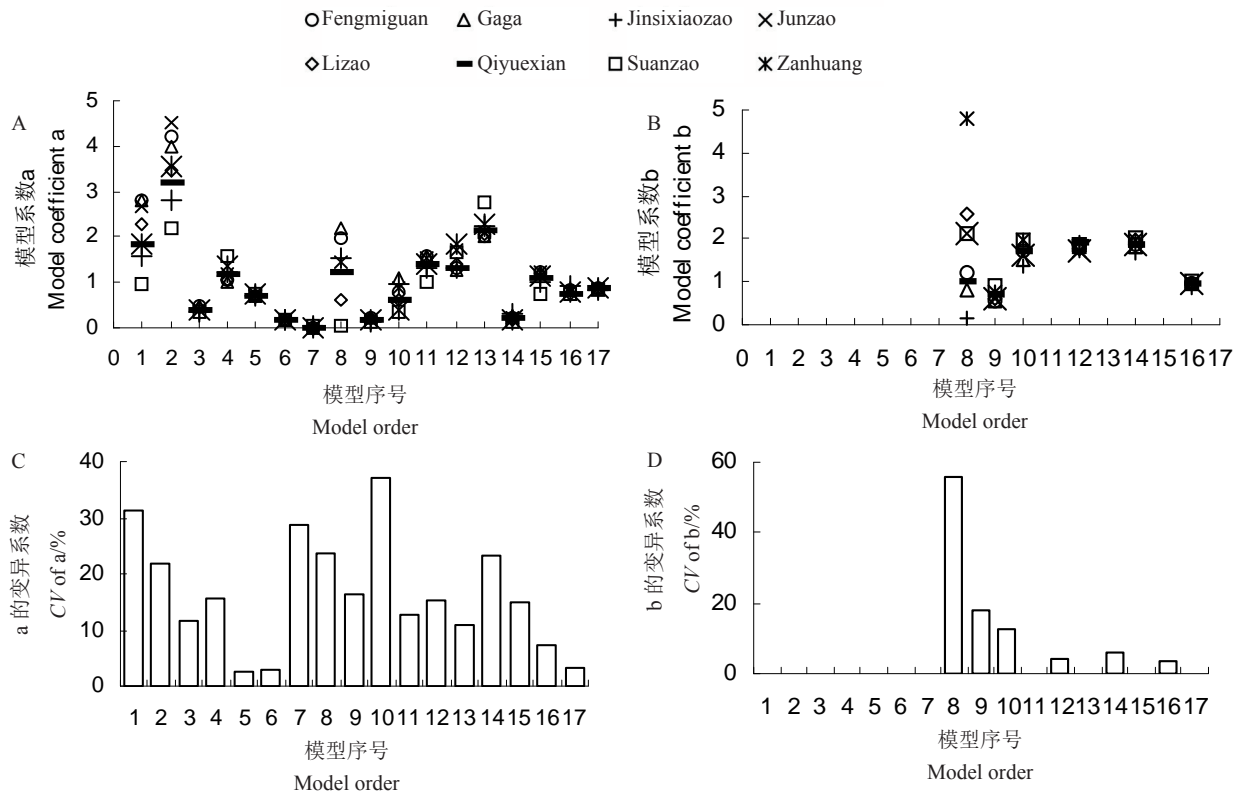


图3 各个品种拟合模型的模型系数 a (A)和模型系数 b (B)值分布及模型系数 a (C) 和模型系数 b (D) 的变异系数(CV)
 Fig.3 The model's fitting coefficient a (A) and coefficient b (b) values and their variation of coefficient (CV) for a (C) and b (D)

6.27%和3.73%，且各个品种模型参数b与不考虑品种差异的总体模型参数b接近。虽然模型12、14和16的参数b具有较小的变异系数，但其模型系数a的变异系数较大，分别为15.32%、25.23%和7.18%。

因此，可以看出，基于模型5、6、9和17，各个品种皆可以准确地模拟各个品种的叶片面积。而模型

5、6和17各个品种参数间的变异系数小，稳定性高，说明这3个模型具有潜在适宜于各个品种的潜力。这3个模型分别为5:LA = a (LL×LW)、模型6:LA = a (LL+LW)²和模型17:LA = (LL × LW) b。

2.4 各个品种对叶面积总体模型的适宜性分析

表5为不考虑品种差异的总体模型5、6和17对

表5 模型5、6和17的总体模型模拟精度与各个品种独立模型模拟精度差异
 Table 5 Differences between models by pooling all data together and models fitted by data of each cultivar for model 5, model 6 and model 17

品种 Cultivar	模型5 Model 5		模型6 Model 6		模型17 Model 17	
	$\Delta R^2/\%$	$\Delta RMSE/cm^2$	$\Delta R^2/\%$	$\Delta RMSE/cm^2$	$\Delta R^2/\%$	$\Delta RMSE/cm^2$
骏枣 Junzao	-0.003 7	-0.047 7	1.022 3	-0.146 0	-0.665 2	0.162 8
嘎嘎枣 Gagazao	-0.280 2	-0.001 4	0.272 2	-0.077 1	0.295 7	-0.056 6
蜂蜜罐 Fengmiguan	0.525 8	-0.238 6	-0.317 0	0.086 8	-0.602 4	0.127 0
梨枣 Lizao	-0.312 0	0.067 4	-0.773 3	0.145 1	-0.661 1	0.077 3
赞皇大枣 Zanhuangdazao	-1.123 2	0.136 6	-0.165 0	0.109 3	-0.253 7	0.040 6
七月鲜 Qiyuexian	-0.019 1	0.014 2	-0.433 3	0.099 5	-2.440 3	0.110 3
金丝小枣 Jinsixiaozao	1.094 7	-0.075 2	0.074 5	-0.037 8	-13.111 4	0.440 5
酸枣 Suanzao	-0.943 2	0.073 2	-1.503 9	0.035 1	-10.703 8	0.177 2

注： ΔR^2 和 $\Delta RMSE$ 分别表示根据所有品种拟合的模型 R^2 和 $RMSE$ 与根据各个品种的数据拟合的 R^2 和 $RMSE$ 间的差值。

Note: ΔR^2 represent the difference between R^2 for the model fitted by pooling all data together and R^2 for the model fitted by data of each cultivar. The same as $\Delta RMSE$.

各个枣品种叶面积的预测精度与各个品种拟合的5、6、17模型的预测精度之间的差值,以验证模型5、6和17的总体模型是否适宜于各个品种。 ΔR^2 为正值,说明总体模型预测的 R^2 值升高。 $\Delta RMSE$ 为负值说明总体模型模拟的RMSE下降,模拟精度提升。从表5可以看出,总体模型5和模型6相比各个品种的独立模型, ΔR^2 最小值出现在模型6的‘酸枣’品种,其总体模型精度($R^2 = 96.80\%$)较其独立模型($R^2=98.30\%$)下降1.503 9%,但依然高于95%。个别品种的总体模型 R^2 值提升,如‘蜂蜜罐’、‘金丝小枣’(模型5)和‘骏枣’、‘嘎嘎枣’和‘金丝小枣’(模型6)。总体模型5和6的RMSE与各个品种的独立模型RMSE的差值 $\Delta RMSE$ 介于 $-0.238 6\sim 0.145 1 \text{ cm}^2$, $\Delta RMSE$ 的最大值出现在模型6的‘梨枣’品种,总体模型RMSE($0.531 7 \text{ cm}^2$)较其独立模型($0.386 6 \text{ cm}^2$)

下降 $0.145 1 \text{ cm}^2$ 。

总体模型17预测各个品种叶面积时,‘金丝小枣’和‘酸枣’的 R^2 值从各品种的独立模型的97.27%和94.22%分别下降13.11%和10.70%,其总体模型的 R^2 值为84.16%和83.51%,RMSE也从独立模型的 $0.307 6 \text{ cm}^2$ 和 $0.245 2 \text{ cm}^2$ 分别升高至总体模型的和 $0.748 1 \text{ cm}^2$ 和 $0.422 4 \text{ cm}^2$,分别增加 $0.440 5$ 和 $0.177 2 \text{ cm}^2$ 。

2.5 总体模型样本数量估计

本研究采集了超过3 000个枣树叶片,通过比较不同随机样本下模型系数的变化,有助于明确模型系数确定所需要的样本数量。从图4可以看出,当样本数量 <200 时,随样本数量的增加,模型5和模型6的模型系数不断波动。当样本数量大于250后,模型系数变异幅度不断变小,基本趋于稳定。

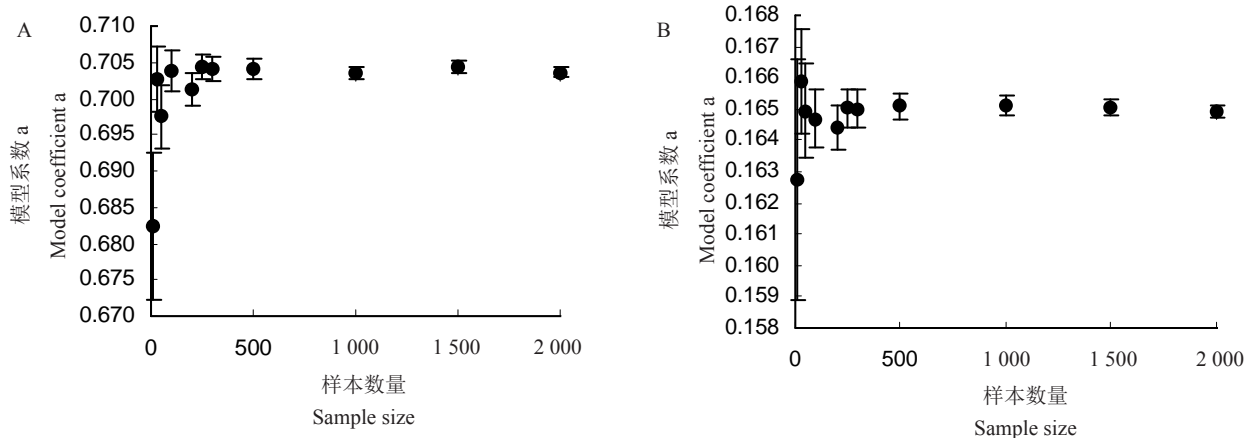


图4 总体模型5(A)和总体模型6(B)的模型系数和标准差随样本数量变化

Fig. 4 Relationships between sample size and model coefficient a and its standard deviation for model 5 (A) and 6 (B) when pooling all data together

3 讨论

基于数学模型的叶片形态参数的叶面积估算已经应用于多种园艺植物。本文对枣树叶面积能否运用基于其叶片形态参数的数学模型进行准确估算进行了系统研究。所选的枣树材料叶片形态参数具有很大的变异性,尤其是叶面积($CV = 64.4\%$),有利于深入了解各个模型的应用潜力及适用范围。

本文中,高精度的模型通常具有较高的 R^2 和较低的RMSE和AIC值。本文总计评价了17个模型的精度,不论总体模型还是各个品种拟合模型,模型精度具有很高的变异性($71.66 \leq R^2 \leq 99.52\%$),说明

能够适用于其他植物的模型,对于枣树而言,并不一定合适^[7,11],需要针对特定植物进行建模。针对枣树的叶面积估算前人亦有报道,但主要针对各个枣品种建模,备选模型较少^[5,17],易忽略潜在的高精度模型,且模型在不同枣品种间的适宜性未有报道。结合文献报道及本文实际,我们总计构建了17个候选模型,通过对采集的数据进行不同处理(1)全数据建模;(2)单个品种数据建模;(3)单个品种对全数据总体模型的适应性。首先,本文筛选出了精度高,适宜于枣叶面积估算的3个模型,即模型5 $LA = a(LL \times LW)$ 、模型6 $LA = a(LL+LW)^2$ 、模型9 $LA = a LL^2 + b LW^2$ 和模型17 $LA = (LL \times LW)^b$ 。模型5与李保国^[17]的研究结果一致,即使用叶长和叶宽的乘积作为变

量,而模型6、9和17则为新发现的潜在高精度模型。对枣叶面积的估算,通常需要针对特定品种建模,获得只能适用于特定品种模型系数^[5,17]。通过对优选模型系数的比较,各品种间模型5和6的系数变异小,且与总体模型系数接近,说明各个品种叶面积的估算可基于相同的模型系数。进一步,通过对各品种适应总体模型的精度与单个品种模型精度的对比,各个品种可使用同样的模型系数估算各个品种的叶面积,可极大的减轻叶面积估算时采集样本的时间和精力。

17个模型中,所有模型自变量为单个叶片形态参数(模型1~4,10~13)或两者的组合(模型5~9,14~17)。对于枣树,当比较以单个叶片形态参数的模型时,以叶宽为变量模型精度要优于同类型以叶长为变量的模型,与咖啡树类似^[20],而与橄榄相反^[13]。一方面可能由于枣叶片宽度(37.34%)的变异系数大于叶片长度(32.20%),对叶面积的贡献可能更大;另一方面,可能由于不同种类植物叶片形态的差异。大量研究认为同时包含叶长和叶宽的叶面积估算模型精度要优于只包含叶长或叶宽的单变量模型^[1,7,13,21],但本文中也发现包含叶长和叶宽的模型 $LA = a(LL+LW)^3$ 和 $LA = aLL + bLW$ 的模拟精度并未优于单变量模型精度,这可能与模型的组织形式以及植物种类有关。本文中筛选的优选模型基于叶长和叶宽两个变量,两个变量的测定时间是单个变量的至少2倍。因此,为提高工作效率,也可利用本文筛选的基于叶宽的单变量模型模型 $LA = aLW^b$ 进行叶面积估算,其 R^2 值大于95%,但对特定品种其模型系数需要进行重新拟合。

多数叶面积模型的构建基于随机采集的叶片,叶片数量并未有明确的限定,如陈宗礼等^[5]使用的叶片数量为40,而李保国^[7]为100~130。本文通过对不同样本量进行建模,发现在叶片数量小于250时,模型系数存在较大的变异,易引起叶面积估算误差的产生。因此,叶面积估算模型的构建需要注意样本数量的控制。

枣树品种众多,叶形态多样,因此,本文筛选出的优选模型是否适宜于其他枣树品种还需要在其他品种枣树进行测试。同时,枣树叶片的生长发育,还受各种环境条件的影响,兼顾基因型和环境的枣树叶面积模型还需要进一步研究。

4 结 论

本文基于8个枣树品种的3287枚叶片形态参数,通过对采集的数据进行全数据建模、单个品种数据建模,以及单个品种对全数据总体模型的适应性分析,构建了适宜本文8个供试枣品种的叶面积估算模型,模型精度可满足叶面积估算需求,进一步明确了枣叶面积估算模型构建的样品数量为>250枚叶片。此研究的开展为枣叶面积的精准估算提供了模型基础,并可辅助三维枣树数字化建模。

参考文献 References:

- [1] KERAMATLOU I, SHARIFANI M, SABOURI H, ALIZADEH M, KAMKAR B. A simple linear model for leaf area estimation in Persian walnut (*Juglans regia* L.)[J]. *Scientia Horticulturae*, 2015, 184: 36-39.
- [2] CRISTOFORI V, ROUPHAEL Y, GYVES E M-D, BIGNAMI C. A simple model for estimating leaf area of hazelnut from linear measurements[J]. *Scientia Horticulturae*, 2007, 113(2): 221-225.
- [3] ROUPHAEL Y, COLLA G, FANASCA S, KARAM F. Leaf area estimation of sunflower leaves from simple linear measurements[J]. *Photosynthetica*, 2007, 45(2): 306-308.
- [4] VAN HOOIJDONK B M, TUSTIN D S, OLIVER M J, DAYATI-LAKE G A. Leaf area development and dry matter production of newly grafted monoaxis and biaxis apple trees[J]. *Acta Horticulturae*, 2016, 1130: 335-340.
- [5] 陈宗礼, 雷婷, 齐向英, 王晓润, 张向前, 刘世鹏, 薛皓. 20个品种枣树叶面积回归方程的建立[J]. *生物学杂志*, 2013, 30(1): 90-94.
CHEN Zongli, LEI Ting, QI Xiangying, WANG Xiaojian, ZHANG Xiangqian, LIU Shipeng, XUE Hao. Established of regression equation of the leaves area on 20 species jujube trees [J]. *Journal of Biology*, 2013, 30(1): 90-94.
- [6] MIGICOVSKY Z, LI M, CHITWOOD D H, MYLES S. Morphometrics reveals complex and heritable apple leaf shapes[J]. *Frontiers in Plant Science*, 2018, 8: 2185.
- [7] DEMIRSOY H. Leaf area estimation in some species of fruit tree by using models as a non-destructive method[J]. *Fruits*, 2009, 64(1): 45-51.
- [8] 杨伟, 高疆生, 徐崇志, 吴翠云, 金强. 红枣叶面积指数与产量的相关性分析[J]. *新疆农业科学*, 2012, 49(8): 1397-1400.
YANG Wei, GAO Jiangsheng, XU Chongzhi, WU Cuiyun, JIN Qiang. The correlation analysis of leaf area index and yield of red jujube[J]. *Xinjiang Agricultural Sciences*, 2012, 49(8): 1397-1400.

- [9] CHEN D, WANG Y, LIU S, WEI X, WANG X. Response of relative sap flow to meteorological factors under different soil moisture conditions in rainfed jujube (*Ziziphus jujuba* Mill.) plantations in semiarid Northwest China[J]. *Agricultural Water Management*, 2014, 136(136): 23-33.
- [10] YANG W, CHEN X, SAUDREAU M, ZHANG X, ZHANG M, LIU H, COSTES E, HAN M. Canopy structure and light interception partitioning among shoots estimated from virtual trees: comparison between apple cultivars grown on different interstocks on the Chinese Loess Plateau[J]. *Trees*, 2016, 30(5): 1723-1734.
- [11] WILLIAMS L, MARTINSON T. Nondestructive leaf area estimation of 'Niagara' and 'DeChaunac' grapevines[J]. *Scientia Horticulturae*, 2003, 98(4): 493-498.
- [12] GYVES E M, ROUPHAEL Y, CRISTOFORI V, MIRA F R J F. A non-destructive, simple and accurate model for estimating the individual leaf area of kiwi (*Actinidia deliciosa*)[J]. *Fruits*, 2007, 62(3): 171-176.
- [13] KOUBOURIS G, BOURANIS D, VOGIATZIS E, NEJAD A R, GIDAY H, TSANIKLIDIS G, LIGOXIGAKIS E K, BLAZAKIS K, KALAITZIS P, FANOURAKIS D. Leaf area estimation by considering leaf dimensions in olive tree[J]. *Scientia Horticulturae*, 2018, 240: 440-445.
- [14] POMPELLI F, ANTUNES C, FERREIRA G, CAVALCANTE S, WANDERLEY-FILHO L, ENDRES L. Allometric models for non-destructive leaf area estimation of *Jatropha curcas*[J]. *Biomass and Bioenergy*, 2012, 36: 77-85.
- [15] 李先明, 秦仲麒, 涂俊凡, 杨夫臣, 朱红艳, 刘先琴. 7个早熟梨品种叶面积回归方程的建立[J]. *江西农业学报*, 2011, 23(5): 36-39.
- LI Xianming, QIN Zhongqi, TU Junfan, YANG Fuchen, ZHU Hongyan, LIU Xianqin. Establishment of regressive equation for leaf area of 7 early-mature pear cultivars[J]. *Acta Agriculturae Jiangxi*, 2011, 23(5): 36-39.
- [16] 刘孟军, 王玖瑞, 刘平, 赵锦, 赵智慧, 代丽, 李宪松, 刘志国. 中国枣生产与科研成就及前沿进展[J]. *园艺学报*, 2015, 42(9): 1683-1698.
- LIU Mengjun, WANG Jiurui, LIU Ping, ZHAO Jin, ZHAO Zhihui, DAI Li, LI Xiansong, LIU Zhiguo. Historical Achievements and frontier advances in the production and research of Chinese Jujube (*Ziziphus jujuba*) in China[J]. *Acta Horticulturae Sinica*, 2015, 42(9): 1683-1698.
- [17] 李保国. 枣树品种叶面积简易测定法[J]. *河北林果研究*, 1989, 4(1/2): 46-50.
- LI Baoguo. A study on a simple method for the determination of leaf area of various jujube trees[J]. *Journal of Hebei Forestry College*, 1989, 4(1/2): 46-50.
- [18] R Development Core Team. R: A language and environment for statistical computing[CP]. ProQuast, 2011.
- [19] SAKAMOTO Y, MASATO I, KITAGAWA G. Akaike information criterion statistics[M]. Netherlands: Springer, 1986: 290.
- [20] ANTUNES W C, POMPELLI M F, CARRETERO D M, DAMATTA F M. Allometric models for non-destructive leaf area estimation in coffee (*Coffea arabica* and *Coffea canephora*) [J]. *Annals of Applied Biology*, 2008, 153(1): 33-40.
- [21] FALLOVO C, CRISTOFORI V, DE-GYVES M, RIVERA M, REA R, FANASCA S, BIGNAMI C, SASSINE Y, ROUPHAEL Y. Leaf area estimation model for small fruits from linear measurements[J]. *Hortscience*, 2008, 43(7): 2263-2267.